(2)

CLONING AND NUCLEOTIDE SEQUENCES OF CROTAMINE GENES

LEONARD A. SMITH* and JAMES J. SCHMIDT

Department of Toxinology, Pathology Division, United States Army Medical Research Institute of Infectious Diseases, Frederick, MD 21701-5011, U.S.A.

Running Title: Crotamine Cloning and Sequencing

*To whom reprint requests should be addressed

L. A. Smith and J. J. Schmidt Cloning and nucleotide sequences of crotamine genes. *Toxicon* --, -----, 19--. A cDNA library containing snake toxin genes was constructed in bacteriophage lambda by using mRNA isolated from the glands of the South American rattlesnake, *Crotalus durissus terrificus*. The first high-density screening of 400,000 plaques for crotamine-containing genes yielded over 800 positives when a labeled cDNA probe with sequence homology to crotamine was used. Four of these clones with insert sizes from 270 to 400 basepairs were chosen and their inserts subcloned into pGEM-3Z and sequenced. Nucleotide sequence analysis of the cloned cDNAs predicted the existence of multiple variants of the crotamine toxin. The different forms, identified from the DNA sequences, displayed discrepancies in amino acid sequence for crotamine when compared with previously published reports. Direct amino acid sequencing of commercially purified crotamine and CNBr fragments thereof confirmed the structures predicted by the nucleic acid sequences.

INTRODUCTION

Animal venoms contain many different pharmacologically active protein components. Many of these constituents are highly toxic to mammals and insects by virtue of their action on nerve and muscle tissues. For example, the venom from the South American rattlesnake, *Crotalus durissus terrificus*, contains at least four distinct groups of protein toxins. These toxins are the presynaptically-acting neurotoxin, crotoxin (SLOTTA and FRAENKEL-CONRAT, 1938; CHANG and LEE, 1977; HAWGOOD and SANTANA DE SA, 1979); a convulsion-producing toxin, convulxin (BRAZIL, 1972; LEE, 1972); a thrombin-like enzyme, gyrotoxin (BARRIO, 1961; ALEXANDER *et al.*, 1988); and a membrane-damaging toxin, crotamine (GONCALVES and POLSON, 1947; GONCALVES and VIEIRA, 1950; GONCALVES, 1961). The latter toxin, crotamine, appears to affect the functioning of voltage-sensitive sodium channels of skeletal muscle sarcolemma, inducing a sodium influx resulting in a depolarization and contraction of skeletal muscle (GONCALVES, 1956; CHEYMOL *et al.*,1971; PELLEGRINI FILHO *et al.*, 1978). In skeletal muscle lesions from the effects of crotamine consist of necrosis of the muscle fibers characterized by extensive vacuolization of the sarcoplasmic reticulum and disruption of actin and myosin filaments (CAMERON and TU, 1978; TU, 1983). Crotamine is a strongly basic polypeptide composed of 42 amino acid residues tightly reticulated by three disulfide bonds (LAURE, 1975; GIGLIO, 1975). The primary sequence of crotamine (LAURE, 1975) exhibits a high degree of homology with other myonecrosis-causing toxins such as peptide C from *Crotalus viridis helleri* venom (MAEDA *et al.*, 1978), myotoxin a from *Crotalus viridis viridis*

venom (FOX *et al.*, 1979), and myotoxins I and II from *Crotalus viridis concolor* venom (BIEBER *et al.*, 1987). In our continuing investigations into the molecular biology of animal toxin genes, recombinant DNA technology has been used to demonstrate the existence of polymorphic variants of crotamine, a protein considered previously to be a single molecular species. We present evidence, obtained from both nucleotide sequences of cloned genes and amino acid sequences of purified crotamine proteins, that the venom from *Crotalus durissus terrificus* contains multiple sequence variants of crotamine.

MATERIALS AND METHODS

Specimens of *Crotalus durissus terrificus* were purchased from Herpetofauna, Fort Myers, Florida. Crotamine, lot number CTM-17, was purchased from Miami Serpentarium Laboratories, Salt Lake City, Utah. All bacterial strains used were Escherichia coli K12 derivatives. Bacterial strain MBM7014 served as host for the primary cDNA library (MEISSNER *et al.*, 1987) and JM109 (YANISCH-PERRON *et al.*, 1985) was the host for propagating the cloned crotamine genes in the pGEM-3Z vector (Promega Biotec, Madison, Wisconsin).

Phage λ packaging extracts were purchased from Strategene (San Diego, California). A synthesis kit for constructing complementary DNA (cDNA) was purchased from Amersham (Arlington Heights, Illinois) and K/RT DNA sequencing system as well as pGEM vectors were purchased from Promega , Biotec. Restriction enzymes were obtained from Pharmacia (Piscataway New Jersey) and DNA methylases from New England Biolabs (Beverly,

Massachusetts). Molecular weight markers for sizing DNA fragments were obtained from BRL, Inc., Gaithersburg, Maryland. Trifluoroacetic acid (TFA), "sequenal grade," was obtained from Pierce Chemical Company, Rockford, Illinois, and 4-vinylpyridine was from Sigma Chemical Co., St. Louis, Montana. The mono-S cation exchange column and related chromatography instruments were obtained from Pharmacia Fine Chemicals, Uppsala, Sweden. Equipment for reverse-phase chromatography was from Waters Associates, Milford, Massachusetts. The column was a Hi-Pore RP-318 from Bio-Rad Laboratories, Richmond, California.

RNA isolation

Glands from three mature *Crotalus durissus terrificus* were removed 3 days after venom extraction (ROTENBERG *et al.*, 1971). The glands were frozen immediately in liquid nitrogen and kept at -70^{0}C until used. Total RNA was isolated from the glands by the guanidinium isothiocyanate-hot phenol extraction method (FERAMISCO *et al.*, 1982). Poly(A)$^+$ RNA was enriched for by passage over an oligo(dT)-cellulose column (AVIV and LEDER, 1972). A yield of 375 µg of total RNA was isolated from 1 g of venom gland tissues. After the total RNA was passed over oligo(dT)-cellulose, 7 ug of poly (A)$^+$ RNA was obtained.

Synthesis and preparation of cDNA for cloning

Complementary DNA was prepared from 2 ug of rattlesnake mRNA (GUBLER and HOFFMAN, 1983) by using a cDNA synthesis kit supplied by Amersham. The synthesized, doublestranded cDNAs were treated subsequently with site-specific BamH I and Alu I methylases to protect internal restriction enzyme sites. After methylation, a bifunctional oligonucleotide linker was ligated to the double-stranded cDNAs. This linker creates a

BamH I restriction site on the 5' end and a Hind III site on the 3' end when ligated to the ends of the cDNAs and restricted with BamH I and Hind III endonucleases (MEISSNER *et al.*, 1987). The modified cDNAs to be cloned were inserted directionally into a λORF8 vector at BamH I and Hind III positions and subsequently packaged into viable phage particles with phage λ packaging extracts.

Analysis of the cDNA library

Library screening was performed according to standard procedures (MANIATIS *et al.*, 1982). The probe used for screening was a nick-translated, 200 basepair (bp) cDNA fragment encoding myotoxin a and its 3' untranslated region (LORIDAN and MIDDLEBROOK, 1988). Hybridization conditions for the myotoxin a cDNA probe were 68^{0}C for 18 hr in 0.09 M sodium citrate/0.90 M NaCl, pH 7.0 (6X SSC) containing 1X Denhardt's solution, 0.1% SDS, and 100 µg/ml denatured salmon sperm DNA. Washing conditions were at 52^{0}C in 0.1X SSC. Autoradiography was carried out at -70^{0}C for 16 h with Kodak X-OMAT XAR5 film.

Subcloning and DNA sequence analysis

Inserts from the four crotamine clones were removed from the λORF8 DNA by a double restriction digest of Hind III and BamH I, gel purified, and ligated into BamH I-and Hind III-cleaved pGEM-3Z DNA. The recombinant plasmids were used to transform competent E. coli JM109 cells. Ampicillin resistant colonies harboring recombinant pGEM-3Z vector constructs were selected by indicator media containing isopropylthio-β-galactosidase (IPTG) and 5-bromo-4-chloro-3-indolyl-β–D–galactosidase (X-Gal) (RUTHER, 1980). Bacterial colonies unable to produce β-galactosidase were

propagated in 10 ml of Luria-Bertani medium containing 50 µg/ml ampicillin, and plasmid DNA was extracted from the cells by the alkaline lysis method (ISH-HOROWICZ and BURKE, 1981). Inserts were evaluated by double restriction enzyme analysis. Four recombinant clones containing cDNA inserts were further propagated in larger quantities and their plasmid DNA purified by CsCl/EtBr density gradient ultracentrifugation (MANIATIS *et al.*, 1982). The DNA sequence of double-stranded plasmid DNA containing the crotamine inserts was analyzed with the Promega K/RT sequencing system. In this method, supercoiled DNA is alkali denatured, primed with either a T7 or SP6 promoter oligonucleotide primer, and sequenced by using dideoxynucleotides and Klenow fragment.

Chemical modifications

For pyridylethylation of cysteines, 2 mg of crotamine was dissolved in 1.0 ml of 6.0M guanidinium hydrochloride/0.50M tris hydrochloride, pH 8.5. Reduction with 0.085M 2-mercaptoethanol was carried out under nitrogen for 6 hr at room temperature. Subsequently, 4-vinylpyridine was added to a final concentration of 0.13M. After 1 hr the mixture was percolated through a C18 Sep-Pak cartridge (Waters Associates, Milford, Massachusetts). After elution of salts and by-products with 0.06% TFA/10% acetonitrile, reduced, pyridylethylated crotamine (RPE-crotamine) was eluted with 0.06% TFA/70% acetonitrile. Cyanogen bromide (CNBr) cleavage of RPE-crotamine was achieved by dissolving 1.4 mg of RPE-crotamine in 0.40 ml of 0.20M cyanogen bromide/70% formic acid (GROSS, 1967). After 24 hr in the dark at room temperature, the mixture was lyophilized. The residue was dissolved in 0.5 ml 0.06% TFA.

Sequence analyses

Automated Edman degradation was done in a model 470A amino acid sequencer from Applied Biosystems, Foster City, California. Phenylthiohydantoin derivatives of amino acids were identified with a model 120A liquid chromatograph from the same manufacturer.

RESULTS

Analysis of mRNA

Primary transcripts were synthesized using polyadenylated messenger RNA (mRNA) purified from the glands of *Crotalus durissus terrificus* and oligo pd(T)$_{12-18}$-primed, reverse transcriptase. Double-stranded cDNA constructs were subsequently prepared with ribonuclease H to remove mRNA while preparing efficient primers for E. coli DNA polymerase I to complete the second-strand synthesis. Although the secondary transcripts ranged in size from 0.2 Kilobasepair (Kb) up to 6.0 Kb, two main bands were observed, with average sizes of 0.40 Kb and 0.75 Kb (Fig. 1). Since crotamine (MW, 4600) and crotoxin (MW, 24,000) are major components of the venom (GONCALVES, 1956; SLOTTA and FRAENKEL-CONRAT, 1938), it is likely that 0.40 Kb and 0.75 Kb could be the expected size for mRNAs coding for crotamine and crotoxin, respectively.

Screening of the cDNA library

Double-stranded cDNAs were inserted into a λORF8 DNA vector and packaged into viable phage particles as described above. After infection of E. coli MBM7014 by the recombinant phage particles, the resulting primary library contained 1.6 X 10^6 plaque-forming units per µg of mRNA

used to construct the library. After amplification of the primary library, 4×10^5 plaques were screened for the crotamine gene by using a radiolabeled probe having sequence homology to crotamine. The first high-density screening of 400,000 plaques yielded 800 positives. Ten clones were selected at random and re-plaqued. After surveying the insert sizes from purified phage DNA using a BamH I and Hind III enzyme double digest, four cDNA inserts were subcloned into pGEM-3Z. Transformants representing each cDNA subcloned were analyzed for inserts by using BamH I and Hind III enzymes. The inserts from subclones pCM45B-8, pCM31B-6, pCM26B-6 and pCM20B-1 had sizes of 400 bp, 350 bp, 340 bp, and 270 bp, respectively, when excised with BamH I and Hind III enzymes (Fig. 2).

Nucleotide sequence of crotamine cDNAs

Figure 3 depicts the nucleotide sequences determined for pCM45B-8, pCM31B-6, pCM26B-6 and pCM20B-1 cDNA. The sequences for pCM45B and pCM26B, aligned to illustrate homology with that of pCM45B, were probably full-length clones beginning with the same sequence, AAGCAG. The nucleotide sequence for pCM45B and pCM26B encode mRNAs with single, continuous, translational reading frames of 208 nucleotides, extending from the BamH I cloning site (GGATCC) through an AUG start codon, to the UAA termination codon. Other features of these crotamine mRNAs include a 5' nontranslated region of 13 nucleotides and a 3' nontranslated region of 116 nucleotides for pCM45B and 109 nucleotides for pCM26B, followed by a polyadenylation tract. A putative polyadenylation signal (AAUAAA) found 26 nucleotides upstream from the beginning of the tract in the case of pCM45B and 19 nucleotides upstream for pCM26B (PROUDFOOT and BROWNLEE, 1976). The crotamine mRNAs predict protoxins of 65 amino

acids which are comprised of a 22 amino-acid signal peptide, a mature crotamine of 42 amino acid residues, and a carboxyl-terminal lysine residue, which is removed during post-translational modification of crotamine. The 5' nontranslated region and the leader sequence for clones pCM45B and pCM26B were identical, coding for 22 residues of predominantly hydrophobic amino acids (Fig. 4). Differences between pCM45B and pCM26B were observed in the structural gene and in the the 3' nontranslated region. The amino acid for pCM45B at position 15 is glutamic acid (GAG) while pCM26B has glycine (GGG) in this position (Fig. 3). The 3' nontranslated sequence for pCM45B and pCM26B is identical except for the additional sequence of (AAACGCT) in pCM45B prior to the polyadenylation tract.

The cDNAs from pCM31B and pCM20B had deletions in the 5' nontranslated regions extending into the start codon for pCM31B and well into the leader sequence of pCM20B (Fig. 4). The region of the genes coding for the native protein were intact in all of the crotamine cDNAs sequenced. Clone pCM20B, although having a large deletion in the 5' end of the gene, had a structural gene that closely resembled the published sequence (LAURE, 1975) except at position 19 (Fig. 8). The clones for pCM45B, pCM31B, pCM26B and pCM20B had leucine rather than isoleucine in this position. The clones pCM45B, pCM31B and pCM26B had arginine in residue 3, isoleucine in residue 6, proline in residue 31 and arginine at residue 34. In pCM20B these residues were glutamine, lysine, arginine, and tryptophan, respectively (Fig. 8). Clone pCM31B had additional modifications in the structural gene at positions 36 and 37. The amino acids normally found at positions 36 and 37 in myotoxin a, peptide C, myotoxin I, and crotamine are cysteine

residues; while in pCM31B, positions 36 and 37 were serine and leucine, respectively.

Amino acid sequence of RPE-crotamine

Residues that could be identified unequivocally from sequencing intact RPE-crotamine are presented in Figure 5. One residue was obtained for each cycle of Edman degradation with the exception of cycle 6, where, in addition to lysine, a small amount of isoleucine was always found. The molar ratio of lysine:isoleucine was approximately 6:1.

Cation exchange chromatography of crotamine

Crotamine was purified further by cation exchange chromatography (Fig. 6). As indicated in the elution profile, selected fractions were pooled, pyridylethylated, and individually placed in the sequencer. In each case, the first 20 residues were identified unequivocally, and were identical to those shown in Figure 5. However, each fraction pool had a different molar ratio of lysine:isoleucine at residue 6. For I, II and III, the ratios were 3:2, 7:1 and 20:1, respectively. Essentially no isoleucine was found at residue 6 in IV.

Purification and sequencing of CNBr peptides

Reverse-phase chromatography of CNBr-cleaved RPE-crotamine is shown in Figure 7a, and the amino acid sequence for fraction pools 1, 3 and 7 are presented in Figure 7b. The sequence of fraction pool 2 was identical to that of 1, while 4, 5 and 6 gave the same results as 3. Sequences of fraction pools 8, 9 and 10 were identical to 7.

DISCUSSION

The nucleic acid sequences reported in this manuscript encode crotamine amino acid sequences with differences from the published structure (LAURE, 1975). Consequently, structural studies of crotamine were performed in order to investigate further these discrepancies. The first study involved automated amino acid sequencing of commercially obtained crotamine (Fig. 5). Identities of the first 20 residues agreed with those based on nucleic acid sequences of clones pCM45B and pCM31B (Fig. 8), with the exception that a mixture of lysine and isoleucine was found for residue 6, instead of isoleucine only. No glutamine was found for residue 3, nor any significant amount of leucine for residue 19. The amino acid sequence also indicated the crotamine purchased commercially had no other interfering contaminating protein species.

The presence of both lysine and isoleucine at residue 6 in crotamine suggested that the commercial preparation was not homogeneous. Therefore, crotamine was purified further to resolve polypeptide chains containing lysine from those with isoleucine at residue 6 (Fig. 6). Fraction pool IV contained only lysine at residue 6, but none of the fraction pools contained sequences with only isoleucine at that location. Although we were unable to purify the isoleucine variant completely, it was clear that about 14-15% of the polypeptide chains present in crotamine have isoleucine at residue 6.

We were not able to identify amino acids beyond residue 20 in the sequence studies described above. In order to obtain structural data in the carboxy-terminal region, crotamine was digested with CNBr and the

resulting fragments were purified by reverse phase chromatography (Fig. 7a). In fraction pools 1 and 2, we found peptides containing the sequence deduced from nucleic acid sequencing: proline instead of arginine at the location corresponding to residue 31 of whole crotamine, and arginine instead of tryptophan for residue 34. Fraction pools 3 through 6 contained the published sequence.

Although fraction pools 1 and 2 produced identical amino acid sequences, their chromatographic behavior was obviously different. The same can be said of pools 3 through 6, and of 7 through 10. We did not determine the basis for this, but possible explanations include the presence of multiple isomers in the 4-vinylpyridine, partial carboxy-terminal amidation, or the presence of residues beyond those identified. Results indicated that our commercial preparation of crotamine contained at least two polypeptides in unequal amounts, a situation that made it impossible to sequence intact crotamine completely. Nonetheless, with the exception of glycine for glutamic acid at residue 15, we showed that amino acid sequences encoded by the nucleic acid sequences are present in the mixture of polypeptides called crotamine.

The nucleotide sequence data of four randomly chosen crotamine genes raise many questions about the diversity and regulation of these genes in crotaline species. Of four genes sequenced, all contained differences in the structural gene, while two of these cDNAs contained deletions in the 5' end and leader sequence of the genes, essentially making them inactive. The pCM20B clone contained a cDNA insert, which, if transcribed and translated, would code for a crotamine protein very similar to the one described by Laure (1975) except for the substitution of isoleucine for leu-

cine at position 19. However, this gene contained a deletion that rendered translation of its mRNA impossible. In support of this result, amino acid sequencing of purified crotamine proteins did not reveal the presence of a crotamine species that pCM20B would produce if the gene was active.

Nonetheless, since only four genes were sequenced, we could not rule out the possibility that other genes may be present in the cDNA library capable of producing the species sequenced by Laure. The clone pCM26B, for example, contains a full-length gene copy that could yield a product identical to one coded for by the clone pCM45B, with the exception of glycine being substituted for glutamic acid at position 15. Amino acid sequencing of purified crotamine proteins, however, did not reveal the presence of glycine in position 15. Thus, we postulate that either the expression of the gene represented in clone pCM26B is down-regulated at the level of translation, or the gene that was present in the three snakes used to prepare the cDNA library was absent in the animals whose venom was used in the commercial purification of crotamine(s). It is also possible that other crotamine variants in the venom were separated chromatographically from the final commercial product. Myotoxin a and crotamine are very basic proteins and both chromatograph very similarly on ion-exchange resins; thus it is unlikely that other variants of crotamine proteins with minor amino acid modifications, described in Figure 8, would have been separated chromatographically from the fraction obtained commercially. Similarly, it seems unlikely that the gene in clone pCM26B would be present in the animals used for our studies but not in the snakes for the a commercial venom source.

We cannot explain why two of the four genes sequenced have deletions

in their genes. The number of crotamine cDNAs sequenced was very low in comparison to the number of crotamine-containing clones obtained, so we cannot conclude anything about the organization of these genes at this time. It is interesting to note that the 5' and 3' untranslated regions in the four cDNAs are relatively conserved while changes in the structural genes have occurred. Also, excision of the cDNA inserts with BamH I and Hind III show a 50 to 130 bp difference in size from the inserts relative to the insert from pCM45B (Fig. 2). This implies that the poly A tracts in the clones have different lengths since this size differential is not accounted for in the cDNA sequences.

In reference to additional post-translational modifications, the removal of the terminal lysine prior to the termination codon (Fig. 3) raises yet another interesting question. About half of the bioactive peptides found in the nervous and endocrine systems of animals have C-terminal residues that are α-amidated (EIPPER and MAINS, 1988). The immediate precursors of α-amidated peptides have structures represented by -X-Gly-Z-Z where X is the C-terminal amino acid acid in the native peptide that is α-amidated, and Z is either lysine or arginine. In the case of crotamine, it is the lysine removed, leaving a terminal glycine; it is not known whether the crotamines are α-amidated on their carboxyl terminus. We are, at present, investigating this possibility as we continue our studies into the organization of these and other toxin genes.

ACKNOWLEDGEMENTS - We thank Young Choi and Clem DeWitt for their excellent technical assistance and Dr. Scott Weinstein for his many thoughtful discussions.

REFERENCES

ALEXANDER, A., GROTHUSEN, J., ZEPEDA, H., and SCHWARTZMAN, R.J. (1988) Gyrotoxin, a toxin from the venom of *Crotalus durissus terrificus*, is a thrombin-like enzyme. *Toxicon* **26**, 953-960.

AVIV, H. and LEDER, P. (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. natl. Acad. Sci. U.S.A.* **69**, 1408-1412.

BARRIO, A. (1961) Gyrotoxin, a new neurotoxin of *Crotalus durissus terrificus* venom. *Acta Physiol. Latinoamericana* **11**, 224.

BIEBER, A.L., MCPARLAND, R.H., and BECKER, R.R. (1987) Amino acid sequences of myotoxins from *Crotalus viridis concolor* venom. *Toxicon* **25**, 677-680.

BRAZIL, O.V. (1972) Neurotoxins from the South American rattlesnake venom. *J. Formosan Med. Assoc.* **71**, 394-400.

CAMERON, D.L. and TU, A.T. (1978) Chemical and functional homology of myotoxin *a* from prairie rattlesnake venom and crotamine from South American rattlesnake venom. *Biochim. biophys. Acta* **532**, 147-154.

CHANG, C.C and LEE, J.D. (1977) Crotoxin, the neurotoxin of South American rattlesnake venom, is a presynaptic toxin acting like β-bungarotoxin. *Naunyn-Schmidebergs Arch. Pharmacol.* **296**, 159-168.

CHEYMOL, J., GONCALVES, J.M., BOURILLET,F., and ROCH-ARVEILLER, M. (1971) Action neuromusculaire comparee de la crotamine et du venin de *Crotalus durissus terrificus* var. *Crotaminicus*-I. *Toxicon* **9**, 279-286.

CHEYMOL, J., GONCALVES, J.M., BOURILLET, F., and ROCH-ARVEILLER, M. (1971) Action neuromusculaire comparee de la crotamine et du venin de *Crotalus durissus terrificus* var. *Crotaminicus*-II. *Toxicon* **9**, 287-289.

EIPPER, B.A. and MAINS, R.E. (1988) Peptide α-amidation. *Ann. Rev. Physiol.* **50**, 333-344.

FERAMISCO, J.R., HELFMAN, D.M., SMART, J.E., BURRIDGE, and THOMAS, G.P. (1982) Coexistence of vinculin-like protein of higher molecular weight in smooth muscle. *J. biol. Chem.* **257**, 11024-11031.

FOX, J.W., ELZINGA, M., and TU, A.T. (1979) Amino acid sequence and disulfide bond assignment of myotoxin *a* isolated from the venom of prairie rattlesnake *(Crotalus viridis viridis)*. *Biochemistry* **18**, 678-684.

GIGLIO, J.R. (1975) Analytical studies on crotamine hydrochloride. *Analyt. Biochem.* **69**, 207-221.

GONCALVES, J.M., and POLSON, A. (1947) The electrophoretic analysis of snake venoms. *Arch. Biochem.* **13**, 253-259.

GONCALVES, J.M. (1956) Purification and properties of crotamine. In: Venoms, Buckley, E.E., Porges, N. (Eds): No.44, pp. 261-273. Washington: Am. Assoc. Adv. Sci..

GONCALVES, J.M. (1961) Estudos sobre crotamine. Thesis Doctorat Medecine, Ribeirao Preto, Sao Paulo.

GONCALVES, J.M. and VIEIRA, L.G. (1950) Estudos sobre veneno de serpentes brasileiras-I. analise electroforetica. *Ann. Acad. Bras. Cienc.* **22**, 141-149.

GROSS, E. (1967) The cyanogen bromide reaction. In: *Methods in Enzymology*, Vol. XI. pp.238-255 (Hirs, C. H. W., Ed.). New York: Academic Press.

GUBLER, U. and HOFFMAN, B.J. (1983) A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263-269.

HAWGOOD, B.J. and SANTANA DE SA, S. (1979) Changes in spontaneous and evoked release of transmitters induced by the crotoxin complex and its component phospholipase A_2 at the frog neuromuscular junction. *Neuroscience (Oxford)* **4**, 293-303.

ISH-HOROWICZ, D. and BURKE, J.F. (1981) Rapid and efficient cosmid vector cloning. *Nucleic Acids Res.* **9**, 2989-2998.

LAURE, C.J. (1975) Die Primarstruktur des crotamins. *Hoppe-Seyler's Z. Physiol. Chem.* **356**, 213-215.

LEE, C.Y. (1972) Chemistry and pharmacology of polypeptide toxins in snake venoms. *A. Rev. Pharmacol.* **12**, 265-286.

LORIDAN, C. and MIDDLEBROOK, J.L. (1988) Cloning and sequencing of the cDNA coding for myotoxin a. *Toxicon* **27**, 60 (Abstract).

MAEDA, N., TAMIYA, N., PATTABHIRAMAN, T.R., and RUSSELL, F.E. (1978). Some chemical properties of the venom of the rattlesnake, *Crotalus viridis helleri*. *Toxicon* **16**, 431-441.

MANIATIS, T., FRITSCH, E.F. and SAMBROOK, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

MEISSNER, P.S., SISK, W.P., and BERMAN, M.L. (1987) Bacteriophage λ cloning system for the construction of directional cDNA libraries. *Proc. natl. Acad. Sci. U.S.A.* **84**, 4171-4175.

PELLEGRINI FILHO, A., VITAL BRAZIL, O., FONTANA, M.D., and LAURE, C.J. (1978) The action of crotamine on skeletal muscle: an electrophysiological study. In: Toxins: Animal, Plant and Microbial. Proceedings of the 5th International Symposium on Animal, Plant and Microbial Toxins, pp. 375-382, (Rosenberg, P., Ed.). London: Pergamon Press.

PROUDFOOT, N. J. and BROWNLEE, G.G. (1976) 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* **263**, 211-214.

ROTENBERG, D., BAMBERGER, E.S., and KOCHVA, E. (1971) Studies on ribonucleic acid synthesis in the venom glands of *Vipera palaestinae (Ophidia, Reptilia)*. *Biochem. J.* **121**, 609-612.

RUTHER, U. (1980) Construction and properties of a new cloning vehicle, allowing direct screening for recombinant plasmids. *Molec. Gen. Genet.* **178**, 475-477.

SLOTTA, K.H. and FRAENKEL-CONRAT, H.L. (1938) Schlangengifte. III. Mitteil. : Reinigung und Krystallisation des Klapperschlangengiftes. *Ber. dtsch. chem. Ges.* **71**, 1076-1081.

TU, A.T. (1983) Local tissue damaging (hemorrhage and myonecrosis) toxins from rattlesnake and other pit viper venoms. *J. Toxicol.-Toxin Rev.* **2**, 205-234.

YANISCH-PERRON, C., VIEIRA, J., and MESSING, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

FIGURE LEGENDS

Figure 1. Characterization by size of double-stranded cDNAs by using polyadenylated mRNA purified from *Crotalus durissus terrificus* glands. One µg of mRNA was first transcribed by 1.6 µg of oligo pd(T)$_{12-18}$ as primer, 1mM dATP, 1mM dGTP, 1mM dTTP, 0.5 mM dCTP containing 20 µC dCTP [^{32} P- α] and 20 units of AMV reverse transcriptase. Second-strand synthesis was continued and completed by using ribonuclease H and DNA polymerease I. An aliquot of [^{32}P] mRNA copies was electrophoresed on a 1.4% agarose gel and autoradiographed for 16 hr at -70^{0}C with X-AR film. The size markers used consisted of [^{32}P] end-labeled λ DNA-Hind III fragments and non-labeled ϕX174 DNA-Hae III fragments.

Figure 2. cDNA inserts from λORF recombinants were subcloned into pGEM-3Z. Transformants were propagated, and plasmids were isolated and analyzed by a double restriction enzyme assay with BamH I and Hind III. Digests were electrophoresed in a 6% polyacrylamide gel and stained in ethidium bromide. Molecular weight markers used were ϕX174-Hae III fragments and BRL 1Kb ladder markers.

Figure 3. Nucleotide sequence of crotamine cDNAs and inferred amino acid sequence. Signal peptide sequences are displayed with a dashed line (---), the 42-amino-acid mature crotamine sequences are underlined by a solid line (___), and a putative polyadenylation sequence, (AATAAA) found 19 to 26 nucleotides upstream from the poly(A) tail, is underlined (___).

Figure 4. Nucleotide sequence of the 5' nontranslated region and leader sequence of cDNA inserts for pCM45B, pCM31B, pCM26B, and pCM20B. Sequences are aligned for maximum homology on that of pCM45B and the BamH I site, [GGATCC], and the codon representing the start of the structural protein [TAT] is represented by brackets. The predicted amino acid sequence is given above the nucleotide sequence.

Figure 5. Partial amino acid sequence of RPE-crotamine.

Figure 6. Cation exchange chromatography of crotamine. Commercially obtained crotamine (500 µg) was redissolved in 0.050 M 2-(N-morpholino)-ethanesulfonic acid (MES), pH 6.5 (Solvent A), and applied to a Pharmacia Mono S cation exchange column (0.5 cm X 5 cm) equilibrated with the same buffer. Solvent B was 0.05 M MES/1.5 M NaCl, pH 6.5. Flow rate was 2.0 ml/min and column effluent was monitored at 280 nm. Fractions pooled are indicated by horizontal bars.

Figure 7a. Reverse-phase chromatography of CNBr-digested RPE-crotamine. The column was a Bio-Rad Hi-Pore RP-318. Solvent A was 0.06% TFA and solvent B was 0.05% TFA /70% acetonitrile. Flow rate was 1.0 ml/min and temperature was 30^{0}C. Column effluent was monitored at 210 nm. Pooled fractions are indicated by horizontal bars. Figure 7b. Amino acid sequences of fraction pools 1 and 3, and partial sequence of fraction pool 7.

Figure 8. Amino acid sequence of crotamine species inferred from nucleotide sequence and from direct amino acid sequence. Previously published amino acid sequences for peptide C, myotoxin a, myotoxin I and crotamine are aligned for maximum homology with the deduced amino acid

sequences from recombinant clones pCM45B, pCM31B, pCM26B and pCM20B and the sequence of purified crotamine from the direct amino acid sequence. Boxes depict regions where heterogeneity exists between variants of myotoxin/crotamine proteins.

Fig. 1

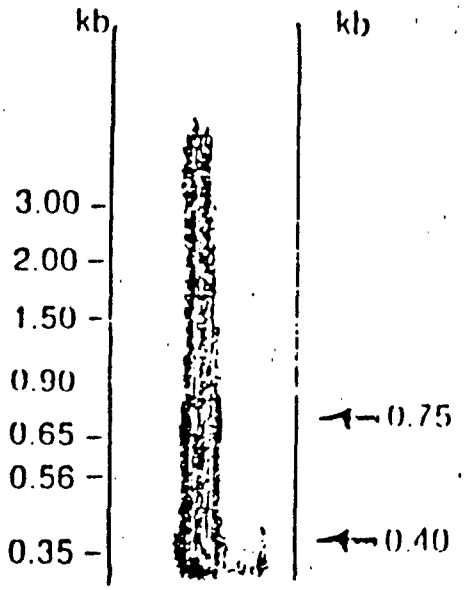


Fig. 2

Fig. 3

```
           10           20       30       40          50        60        70
AAGCAGTCTCAGC ATG AAG ATC CTT TAT CTG CTG TTC GCA TTT CTT TTC CTT GCA TTC CTG TCT GAA CCA GGG AAT
               M   K   I   L   I   L   L   F   A   F   L   F   L   A   F   L   S   E   P   G   N
AAGC            G AAG ATC CTT TAT CTG CTG TTC GCA TTT CTT TTC CTT GCA TTC CTG TCT GAA CCA GGG AAT
                   K   I   L   I   L   L   F   A   F   L   F   L   A   F   L   S   E   P   G   N
AAGCAGTCTCAGC ATG AAG ATC CTT TAT CTG CTG TTC GCA TTT CTT TTC CTT GCA TTC CTG TCT GAA CCA GGG AAT
               M   K   I   L   I   L   L   F   A   F   L   F   L   A   F   L   S   E   P   G   N
AAGC                                                                CA TTC CTG TCT GAA CCA GGG AAT
                                                                        F   L   S   E   P   G   N
             -22                                         -10

 80         90          100           110       120        130         140       150
GCC TAT AAA CGG TGT CAT ATA AAA GGA GGA CAC TGC TTT CCC AAG GAG AAA ATA TGT ATT CCT CCA TCT TCT GAC
 A   T   K   R   C   H   I   K   G   G   H   C   F   P   K   E   K   I   C   I   P   P   S   S   D
GCC TAT AAA CGG TGT CAT ATA AAA GGA GGA CAC TGC TTT CCC AAG GAG AAA ATA TGT ATT CCT CCA TCT TCT GAC
 A   T   K   R   C   H   I   K   G   G   H   C   F   P   K   E   K   I   C   I   P   P   S   S   D
GCC TAT AAA CGG TGT CAT ATA AAA GGA GGA CAC TGC TTT CCC AAG GGG AAA ATA TGT ATT CCT CCA TCT TCT GAC
 A   T   K   R   C   H   I   K   G   G   H   C   F   P   K   G   K   I   C   I   P   P   S   S   D
GCC TAT AAA CAG TGT CAT AAG AAA GGA GGA CAC TGC TTT CCC AAG GAG AAA ATA TGT ATT CCT CCA TCT TCT GAC
 A   T   K   Q   C   H   K   K   G   G   H   C   F   P   K   E   K   I   C   I   P   P   S   S   D
    +1                                  +10                                    +20

      160           170         180         190             200        210      220
TTT GGG AAG ATG GAC TGT CCA TGG AGA CGG AAA TGC TGT AAA AAG GGA AGT GGA AAA TAA TGCCATCTCCATCTAGG
 F   G   K   M   D   C   P   W   R   R   K   C   C   K   K   G   S   G   K  END
TTT GGG AAG ATG GAC TGT CCA TGG AGA CGG AAA TCG CTG AAA AAG GGA AGT GGC AAA TAA TGCCATCTCCATCTAGG
 F   G   K   M   D   C   P   W   R   R   K   S   L   K   K   G   S   G   K  END
TTT GGG AAG ATG GAC TGT CCA TGG AGA CGG AAA TGC TGT AAA AAG GGA AGT GGA AAA TAA TGCCATCTCCATCTAGG
 F   G   K   M   D   C   P   W   R   R   K   C   C   K   K   G   S   G   K  END
TTT GGG AAG ATG GAC TGT CGA TGG AGA TGG AAA TGC TGT AAA AAG AGA AGT GGA AAA TAA TGCCATCTCCATCTAGG
 F   G   K   M   D   C   R   W   R   W   K   C   C   K   K   R   S   G   K  END
                    +30                                   +40

230        240         250          260        270        280        290       300        310
ACCATGGATATCTTCAAGATATGGCCAAGG ACCTGAGAGTGCCGCCTGCTATCGCTTTATCTTTCTTTATCTAAATAAAATTGCT
ACCATGGATATCTTCAAGATATGGCCAAGG ACCTGAGAGTGCCGCCTGCTATCGCTTTATCTTTCTTTATCTAAATAAAATTGCT
ACCATGGATATCTTCAAGATATGGCCAAGG ACCTGAGAGTGCCGCCTGCTATCGCTTTATCTTTCTTTATCTAAATAAAATTGCT
ACCATGGATATCTTCAAGATATGGCCAAGG ACCTGAGAGTGCCGCCTGCTATCGCTTTATCTTTCTTTATCTAAATAAAATTGCT

      320
ACCTATCAAACGCTAAAAAAAA >>        pCM45B
ACCTATCAAACGCTAAAAAAAA >>        pCM31B
ACCTATC-------AAAAAAAA >>        pCM26B
ACCTATC-------AAAAAAAA >>        pCM20B
```

Fig. 4

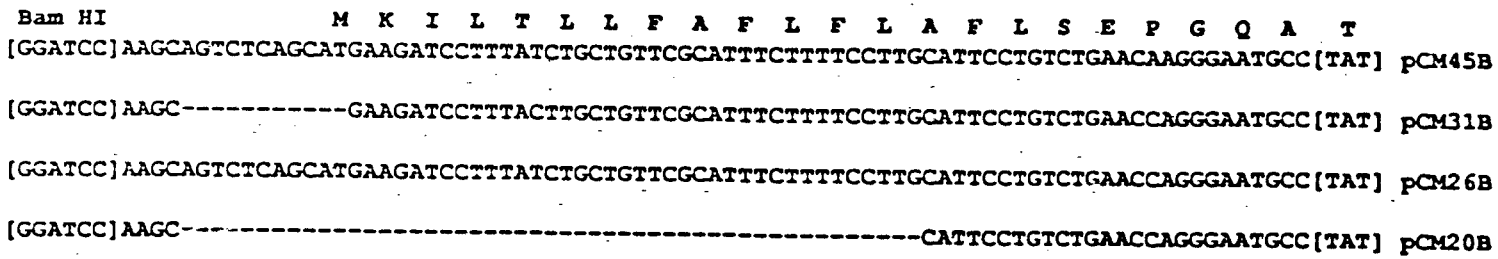

Bam HI
M K I L T L L F A F L F L A F L S E P G Q A T
[GGATCC]AAGCAGTCTCAGCATGAAGATCCTTTATCTGCTGTTCGCATTTCTTTTCCTTGCATTCCTGTCTGAACAAGGGAATGCC[TAT] pCM45B
[GGATCC]AAGC----------GAAGATCCTTTACTTGCTGTTCGCATTTCTTTTCCTTGCATTCCTGTCTGAACCAGGGAATGCC[TAT] pCM31B
[GGATCC]AAGCAGTCTCAGCATGAAGATCCTTTATCTGCTGTTCGCATTTCTTTTCCTTGCATTCCTGTCTGAACCAGGGAATGCC[TAT] pCM26B
[GGATCC]AAGC----------CATTCCTGTCTGAACCAGGGAATGCC[TAT] pCM20B

Fig. 5

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Y	K	R	C	H	K + I	K	G	G	H	C	F	P	K	E	K	I	C	I	P

Fig. 6

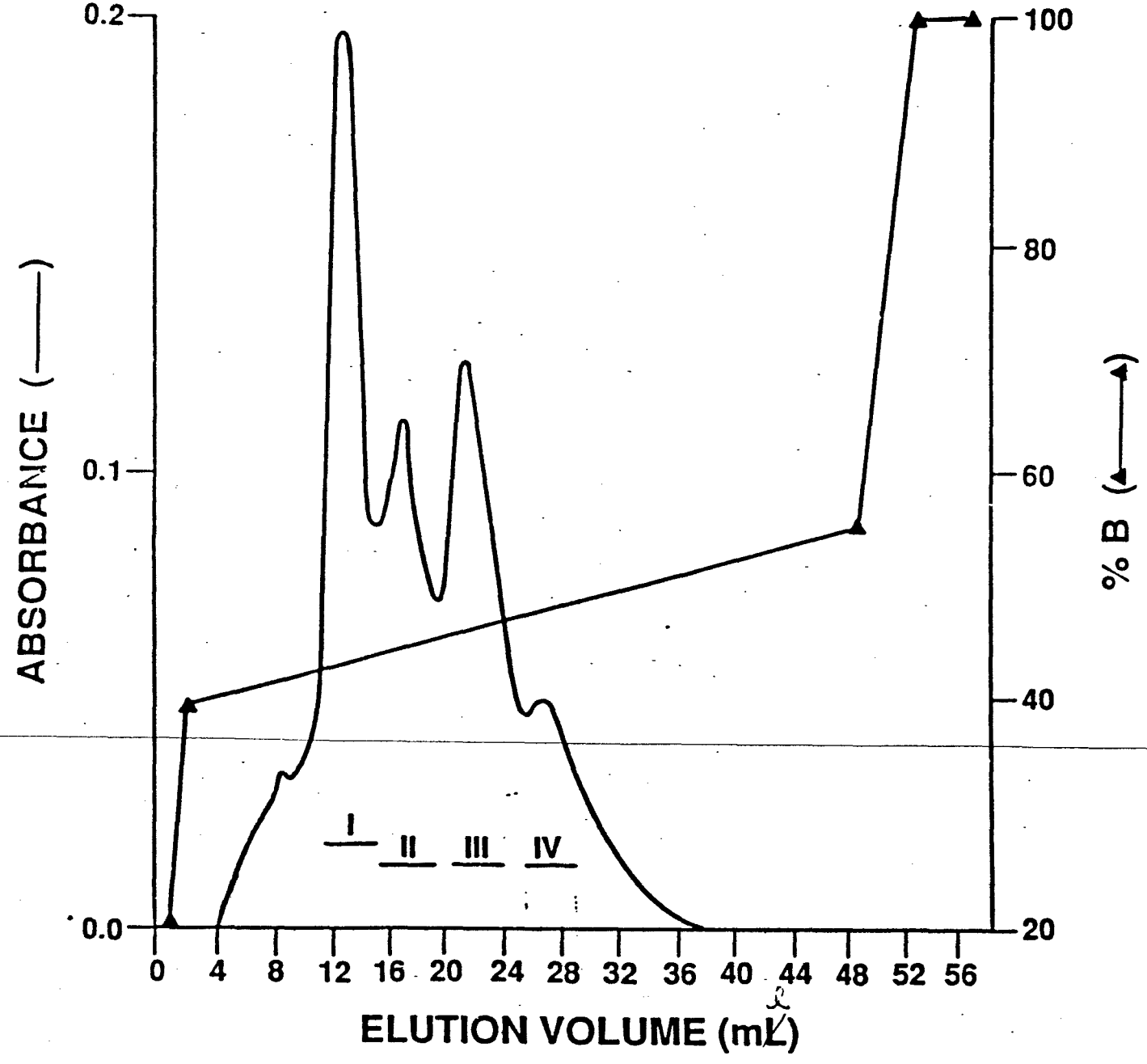
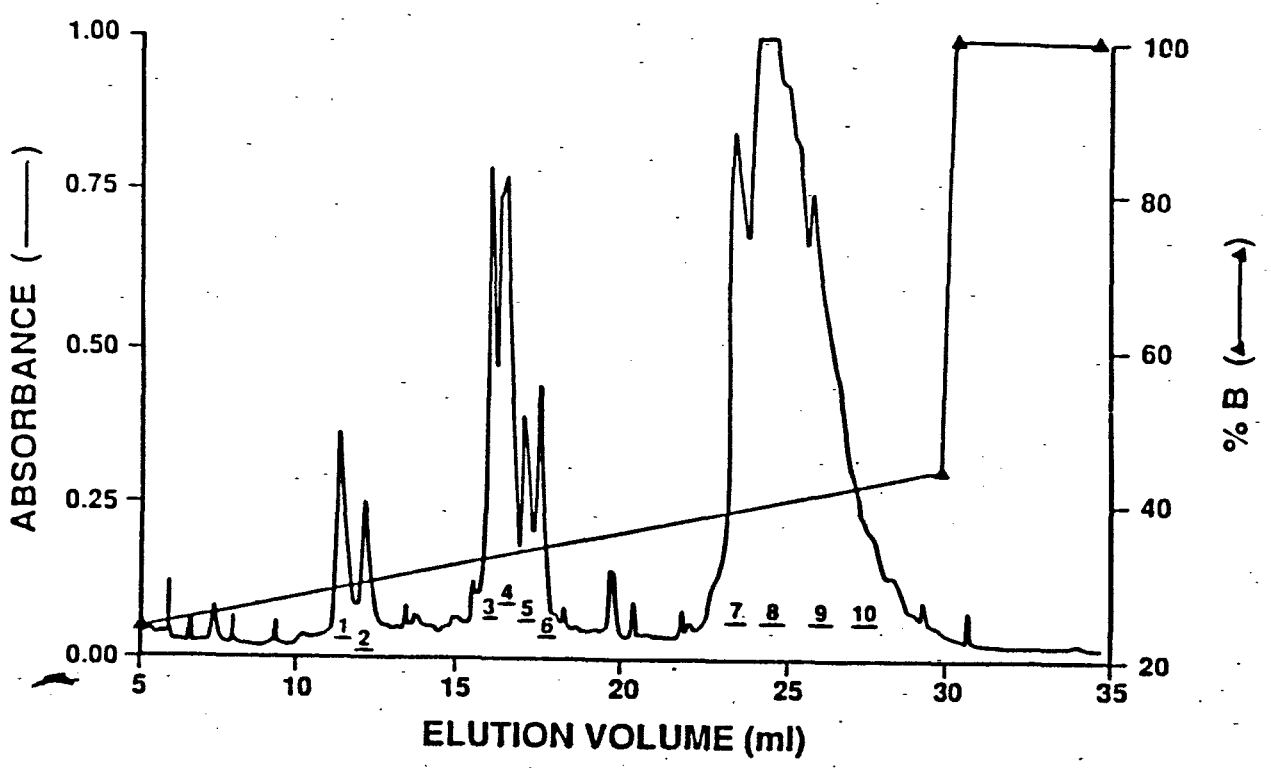

0.2
0.1
0.0
ABSORBANCE (——)
100
80
60
40
20
% B (- - -)
I
II
III
IV
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56
ELUTION VOLUME (mL)

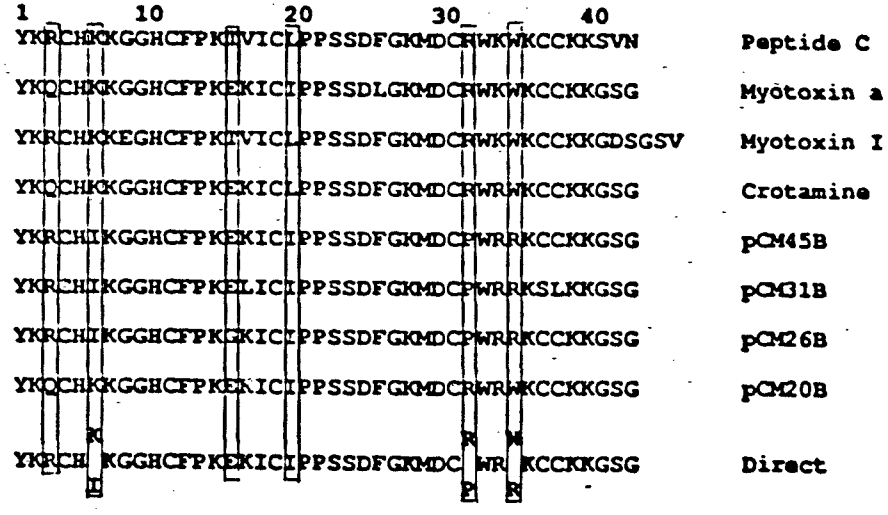
Fig. 7 a and

a



b

Fraction Pool	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	D	C	P	W	R	R	K	C	C	K	K	G	S	G						
3	D	C	R	W	R	W	K	C	C	K	K	G	S	G						
7	Y	K	R	C	H	K + I	K	G	G	H	C	F	P	K	E	K	I	C	I	P


1
10
20
30
40
YKRCHKKGGHCFPKDVICLPPSSDFGKMDCRWKWKCCKKSVN
Peptide C
YKQCHKKGGHCFPKEKICIPPSSDLGKMDCRWKWKCCKKGSG
Myotoxin a
YKRCHKKEGHCFPKDVICLPPSSDFGKMDCRWKWKCCKKGDSGSV
Myotoxin I
YKQCHKKGGHCFPKEKICLPPSSDFGKMDCRWRWKCCKKGSG
Crotamine
YKRCHIKGGHCFPKEKICIPPSSDFGKMDCRWRRKCCKKGSG
pCM45B
YKRCHIKGGHCFPKELICIPPSSDFGKMDCRWRRKSLKKGSG
pCM31B
YKRCHIKGGHCFPKGKICIPPSSDFGKMDCRWRRKCCKKGSG
pCM26B
YKQCHKKGGHCFPKEKICIPPSSDFGKMDCRWRWKCCKKGSG
pCM20B
YKRCH KGGHCFPKEKICIPPSSDFGKMDC WR KCCKKGSG
Direct

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Distribution unlimited - approved for public release
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION U.S. Army Medical Research Institute of Infectious Diseases 6b. OFFICE SYMBOL (If applicable) SGRD-UIP-B	7a. NAME OF MONITORING ORGANIZATION U.S. Army Medical Research and Development Command
6c. ADDRESS (City, State, and ZIP Code) Ft. Detrick Frederick, MD 21701-5011	7b. ADDRESS (City, State, and ZIP Code) Ft. Detrick, Frederick, MD 21701-5012
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Cloning and Nucleotide Sequence of Cromtamine Genes

12. PERSONAL AUTHOR(S)
Drs. Leonard A. Smith and James J. Schmidt

13a. TYPE OF REPORT	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
		24 MAY 89	31

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Cloning and nucleotide sequences of crotamine genes. A cDNA library containing snake toxin genes was constructed in bacteriophage lambda by using mRNA isolated from the glands of the South American rattlesnake, Crotalus durissus terrificus. The first high-density screening of 400,000 plaques for crotamine-containing genes yielded over 800 positives when a labeled cDNA probe with sequence homology to crotamine was used. Four of these clones with insert sizes from 270 to 400 basepairs were chosen and their inserts subcloned into pGEM-3Z and sequenced. Nucleotide sequence analysis of the cloned cDNAs predicted the existence of multiple variants of the crotamine toxin. The different forms, identified from the DNA sequences, displayed discrepancies in amino acid sequence for crotamine when compared with previously published reports. Direct amino acid sequencing of commerically purified cortamine and CNBr fragments thereof confirmed the structures predicted by the nucleic acid sequences.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE